Needs-Based Analysis of Online Customer Reviews

Thomas Y. Lee
University of Pennsylvania, The Wharton School
573 JMHH, 3730 Walnut Street
Philadelphia, PA 19104
1-215-893-9528
thomasyl@wharton.upenn.edu

ABSTRACT
Needs-based analysis lies at the intersection of product marketing and new product development. It is the study of why consumers purchase and what they do with those purchases. In a world of mass-customization and one-to-one marketing, anticipating the customer's needs is a key competitive advantage. In this paper, we consider a new approach to supplement traditional methods for assessing rapidly changing user needs. We model the knowledgebase of online customer reviews as a matrix of reviews relating customer needs to product attributes. In a hierarchical two-stage process, we first use association rules to cluster related attributes and needs into hyper-edges. In a second application of association rule mining, we search for hyper-rules relating hyperedges. The method is demonstrated on 10,500 customer reviews over two unrelated product domains, digital cameras and vacuum cleaners.

Categories and Subject Descriptors
H3.1 [Content Analysis and Indexing]: Abstracting methods. H.4.2 [Types of Systems]: Decision support. H.2.4 [Systems]: Textual databases H.2.8 [Database Applications]: Data mining

General Terms
Management, Design, Experimentation, Human Factors.

Keywords
User needs; customer needs; text mining; customer reviews.

1. INTRODUCTION
Customer reviews are a venue where individuals may share their experiences with and reactions to goods and services ranging from consumer electronics to home interior design. Customers have long formed independent fan clubs around a particular product or practice (e.g. Harley Davidson's Owner's Group [12]. However, the Internet has magnified the influence and scope of independent feedback forums by fostering an entire industry in user reviews.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ICEC'07, August 19–22, 2007, Minneapolis, Minnesota, USA.
Copyright 2007 ACM 978-1-59593-700-1/07/0008...$5.00.

Psychologists and computer scientists have conducted fine-grained analysis, analyzing the text content of customer comments and deconstructing them into a series of positive and negative comments about different product characteristics (e.g. what is good and what is bad about a product or service) [13, 20, 33]. Marketers and economists have documented the impact of "the voice of the consumer" on sales. In reviews, both what is expressed [5] and by whom [30, 37] appear to affect conversion rates. Manufacturers and retailers therefore face a strategic decision on whether to share reviews and when to do so [3].

Prior analysis of online reviews, however, appear to have overlooked the role of *needs*. A "*need*" is defined in the Oxford English Dictionary as "[a] feeling of want that provides a basis for behavior or action [22]." The goal of commerce is to satisfy needs through the transaction of goods and services [14]. Manufacturers and retailers have long gathered information to identify needs by observing and interacting with customers through interviews, focus groups, and post-sales product registrations, rebates, and satisfaction surveys [14, 34]. The emergence of widely-accessible, freely-available, online reviews offers an opportunity to complement existing focus group and survey analysis with data from online reviews.

In this on-going work, we propose to combine linguistic and statistical text processing techniques with more traditional association-rule mining for the purpose of automatically identifying user needs. From an applications perspective, this is the first analysis that we know of to pursue a general approach for eliciting customer needs. Needs-analysis complements managerial judgment and existing survey techniques for bringing the "voice of the customer" to the initial stages of product marketing and new product development. From a technical perspective, we model customer reviews as an implicit bi-partite structure that separates needs from attributes. To discover this implicit structure, we use a multi-level association rule mining approach [10, 11]. We first construct hyper-edges that cluster related terms [10]; a second mining stage then searches for hyper-rules that link hyper-edges.

In the remainder of this paper, we begin by reviewing the needs for needs-analysis that motivate this work. Section 3 is a review of the prior literature applying text processing to online reviews and the relationship to needs-analysis. In section 4, we propose our model for extracting needs from customer reviews; preliminary evaluation results appear in section 5. We revisit the need for needs in our discussion and conclusion.

2. MOTIVATION: THE NEED FOR NEEDS

The fields of product marketing and new product development intersect where customer needs are defined. The generic, needs-driven process depicted in Figure 1 illustrates the role that needs play in both product marketing and new product development. As described in standard marketing texts [14], the end result of the classical purchasing model is Figure 1 Step 4, the customer's purchase of a specific product. However, the process begins with Figure 1 Step 1, the recognition of a particular need. Because many customers may be quite familiar with their own needs yet unfamiliar with technical specifications for selecting among competing products, recognition is followed by Figure 1 Step 2, research. The voice of the consumer, as expressed by needs, is translated into the language of the producer, as represented by product attributes. Once needs are translated to attributes, Figure 1 Step 3 transitions from research to evaluation. Which product (bundle of attributes) will best satisfy a particular need? The process terminates with the purchase decision in Step 4.

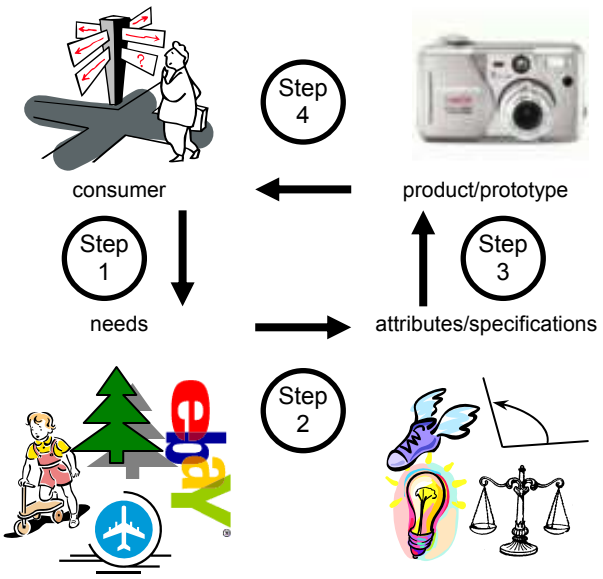


Figure 1. A generic, needs-driven process.

Advances in mass customization and one-to-one marketing [14, 19] demonstrate the relationship between marketing and product development that revolves around understanding customer needs. Phase one of the model product development process is Concept Development [34]. As with the classical purchasing model, the steps of concept development map onto the generic needs-drive process depicted in Figure 1. Step 1 of concept development acknowledges the customer's needs. Step 2 of concept development translates customer needs into "metrics" or technical specifications. In the language of marketing, technical specifications are simply product attributes. In step 3, designers generate and evaluate product concepts where purchasers would evaluate existing products in the marketplace. Finally, product designers would select a set of final specifications (a final concept) just as consumers would select a product purchase.

The challenge for effective product marketing and conceptual design lies in Figure 1 Step 2, speaking to specific customer needs on the customer's terms [14]. Because they cannot directly know the mind of the consumer (Step 1), marketers and designers use identical techniques to learn needs. Consumers historically relied on personal relationships such as trusted family and friends for product research. Today, consumers extend their research to include the community of users represented by various online forums [30, 37]. The objective of this research is to likewise extend the marketing and design understanding of consumer needs to include feedback from these same online forums.

3. RELATED WORK

Our needs-based approach to online reviews lies at the intersection of three different research streams in the literature: Content analysis of online product reviews, Needs-based recommenders, and User-oriented product design.

3.1 Content Analysis of Online Product Reviews

Content analysis has a long history in social science research [25, 36]. However, we began with the automated analysis of sentiment in online customer reviews. Turney [33] uses part-of-speech (POS) analysis to extract sentiment terms from noun phrases. The overall tenor of a review as positive or negative is measured by computing the pairwise mutual information (PMI or co-occurrence frequency) between extracted terms and known sentiment terms in a reference corpus.

Rather than estimate the overall tenor of a review, the follow-on work pairs sentiment terms with specific product attributes. Nasukawa and Yi [20] observed that in reviews, sentiment terms are typically grammatical modifiers of a subject noun representing a product attribute. They combine POS analysis with shallow text parsing to associate sentiment terms with the specific product attributes. Hu and Liu [13] incorporate frequent item-set analysis to accommodate multi-term product attributes (e.g. digital zoom) as well as to discover less-frequently mentioned product features that would fail a co-occurrence test. Popescu and Etzioni [24] extend the use of PMI [33] to more accurately extract product attribute terms paired with sentiments [20] and to cluster synonymous or "implicit" references to a common attribute. Using term frequencies and co-occurrences is generalized in RedOpal, which discovers product attributes by finding noun phrases that deviate from the word distributions of a general, reference corpus [28].

Leveraging WordNet linguistic relations, Popescu and Etzioni further decompose attributes into their corresponding properties [Popescu]. Rather than relying upon a WordNet ontology, Lee [16] uses a constrained-logic program to simultaneously cluster phrases referring to the same attribute and to discover the underlying properties of a given product attribute.

Unfortunately, while existing approaches have successfully addressed product attributes, none have addressed the issue of user needs. Moreover, focusing exclusively on product attributes may inadvertently conflate needs with attributes rather than relating the two. For example, in reviewing barbecue grills, might we regard "hamburger" as a need or attribute [28]? This work complements earlier work on customer reviews by introducing customer needs to the analysis.

3.2 Needs-based Recommenders

The long history of research in recommender systems has focused on product and consumer attributes and transaction histories [1]. The sparing use of customer reviews tended to focus on numerical scores within the reviews [4]. Of the research on customer reviews, researchers initially framed their analysis of customer reviews as efforts at text summarization [4, 13, 20, 33]. However, text summarization has quickly given way to customer decision support; separating feature-specific sentiments enables finer-grained product comparsions [7, 17, 24]. Rather than scoring based upon sentiment, RedOpal [28] joins extracted features with explicit review ratings to numerically score features and products. Ghose and Iperitos numerically score individual features by constructing a hedonic pricing model [8].

Conceptually, the value of a needs-based approach is understood; "Web-assisted needs-based selling should build on information specific to individual customers and their anticipated use of the product [9]." However, little of the prior work on customer review-based recommendation extends back to Step 1, the needs of the consumer.

The prior work on needs-based approaches provides little insight into the necessary knowledge acquisition. O'Keefe and McEachern [21] describe customized applications by GE Plastics and United Technologies, Inc. to enable customers to configure models or specify solution characteristics; but both efforts required millions of dollars in employee-specific knowledge-intensive initial investment. Stolze and Stroble [31] propose a needs-based digital camera recommender using multi-attribute utility theory to score products relative to a consumer's attribute preferences (Step 3 in the general needs-based model of Figure 1). However, they omit a discussion of the critical acquisition of needs (Step 1) and the relationship between needs and the attributes used to generate a recommendation (Step 2).

Orman [23] suggests online communities, transaction histories, manufacturer information, and consumer preferences as contributing sources and raises the research challenge of extracting and integrating needs-based knowledge for customer decision support. AutoChoiceAdvisor, a needs-based vehicle recommender, exemplifies Orman's architecture by integration customer purchase data and product manufacturing details as well as survey results from a third-party, automotive marketing services company [35]. Lee [15] focuses on online customer reviews as a knowledge base for needs-based analysis. Reviews are transformed into Boolean vectors of needs and attributes that are mined for association rules relating needs to attributes. Unfortunately, the ontology-based vector-transformation requires prior knowledge of both needs and attributes. This work contributes to the conceptual pieces on needs-based customer decision support by suggesting one approach for automatically eliciting needs, complementing traditional market research.

3.3 User-oriented Product Design

There is a long history of research in product design [34], in mass customization [14], and in techniques to select attributes combinations that maximize utility [32]. However, there is little work on using automated methods to augment product design. The exception is in mass customization. Randall et al. [26] develop a complementary customer support tool for configuring Dell laptop computers. Users prioritize a set of pre-specified needs such as "video editing" or "email" as inputs to a linear multiattribute utility model that scores all possible attribute configurations. However, identifying the pre-specified needs and constructing the utility function to accurately relate needs to features, requires domain expertise.

Urban and Hauser [35] propose to gather user needs by mining user transactions and interactive chat logs from AutoChoiceAdvisor. In this way, a customer's search for an automobile is used to inform the design of new automobiles. However, in order to be useful for product design purposes, the possible navigation paths must be carefully designed so that engineers can translate customer feedback into specifications [35]. Needs-based analysis of online customer reviews may facilitate the design of interactive systems like AutoChoiceAdvisor by highlighting previously unanticipated needs or relationships.

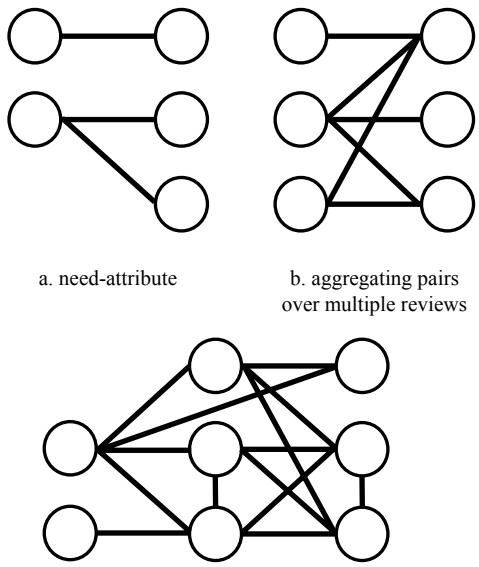
4. APPROACH

The traditional approach to analyzing product reviews models each review as a set of statements about product attributes. Consequently, regardless of method, extracted terms are necessarily interpreted as product attributes (or attribute properties). Our goal is to instead differentiate between needs, as expressed by consumers and actionable features, useful to product designers.

4.1 Intuition

Our basic framework assumes that a review embeds a need-attribute pair (Figure 2a). Of course a single need may relate to multiple attributes and vice-versa. For example, a digital camera for travel may require both compactness and a long battery life. As a product attribute, compactness may support users interested in photos while traveling or in school. In principle, aggregating over reviews would reveal a bipartite structure differentiating needs from features (Figure 2b) as well as hinting at how the two relate.

Unfortunately, knowing only the co-occurrences of different need-words or attribute-words within a review reveals an inherently noisy picture. A single review may include multiple need-attribute pairs or no pairs at all. Reviews with multiple need-attribute pairs lead to ambiguity in knowing which attributes support which needs. Reviews that list attributes or needs exclusively exacerbate the problem of separating needs from attributes. Looking at the co-occurrences of different words for needs and attributes, the resulting dense structure obscures the intuitive bipartite classification of needs and attributes (Figure 2c). Our goal is to introduce a layer of abstraction. By selectively clustering terms together, we aim to dampen the noise and highlight the desired separation (Figure 3).



a. need-attribute

b. aggregating pairs over multiple reviews

c. noisy aggregation of need-attribute

Figure 2. Reviews as graphs relating needs and

4.2 Model

Formally, we model a corpus of customer reviews for a single product category in a Boolean review-token matrix *M* where tokens represent words or phrases that name a user need or product attribute. Note that multiple occurrences of a particular token in a single review are counted only once. The matrix *M* encodes a graph *G* = (*V*,*E*). The set of vertices *V* represents tokens. An edge in *E* indicates the co-occurrence of two vertices within a single review. Edges are weighted by the frequency with which the corresponding pair occurs throughout the corpus. Our goal is then to find those edges that separate needs from attributes.

To dampen the noise in the graph from reviews listing multiple attributes and/or needs, we propose a two-stage, repeated application of association rule mining. We first pre-cluster vertices into a hypergraph *H*= (*V*,*E*) where vertices continue to represent tokens but hyperedges *E* can connect (contain) two or more vertices [29]. As in association rule clustering [10], a hyperedge is defined by the average confidence of all association rules constructed using *all* members (no subsets) of the hyperedge. In a second application of rule mining, we then search for relationships between hyperedges.

The hyper-rule $E \rightarrow F$ defines an association from vertices in the hyperedge *E* to the hyperedge *F*. The support and confidence of a hyper-rule is calculated directly from the co-occurrences on the underlying data set. In traditional association rule mining, if *R* is the set of all reviews, *S* is a set of vertices, and *T* is the set of all reviews containing *S*

$$Support(S) = |T| / |R|$$

The confidence for a rule is traditionally written in terms of

$$Confidence(A \rightarrow B) = Support(AB) / Support(A)$$

support. For hyper-rules, the Support of a single hyperedge (a set of vertices) remains the same. However, for a rule $E \rightarrow F$ relating hyper-edges, a review supports the rule if that review contains at least one vertice (one token) in *E* and one in *F*. For the vertices in the sets *E* and *F*, this suggests that |*EF*| is too restrictive. Instead, we allow for all permutations of a subset of *E* relating to a subset of *F*. The numerator is then:

$|2^E \times 2^F|$ and the denominator remains *Support*(*E*).

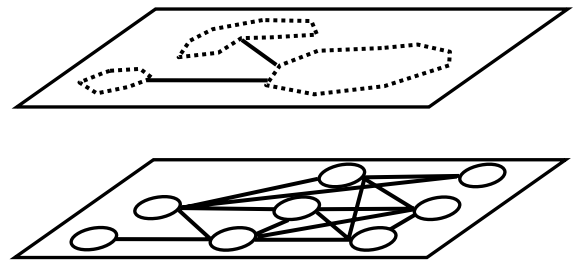


Figure 3. Abstracting the noisy aggregation of need-attribute pairs from Figure 2c.

4.3 Implementation

Beginning with a corpus of customer reviews, we first extract needs and features. Following much of the prior literature on sentiment analysis, we use the Monty Lingua [18] toolkit to parse and chunk online review text into their respective POS phrases. A review sentence such as: "We are expecting our first baby" is transformed to: (NX We/PRP NX) (VX are/VBP expecting/VBG VX) (NX our/PRP$ first/JJ baby/NN NX) ./. We focus on noun phrases although verb phrases are also known to provide resolution on product attributes [33].

A standard stopword list eliminates common pronouns (e.g. "We") and the MontyLingua lemmatiser partially accounts for inconsistencies in tense, plurality, etc. The resulting word and phrase tokens are pruned by eliminating the most and least frequently occurring. A review-token matrix maps token occurrences to individual reviews.

For the two-stage association rule mining, we use the Borgelt implementation of the apriori association rule mining program [2]. In the first stage of mining, we generate association hyper-edges by clustering frequent token-sets. For each cluster, we update the review-token matrix by introducing a new token that appears in every review containing at least one cluster element. In the second stage, we search for association rules in the updated review-token matrix.

5. EVALUATION

We repeated our analysis on two collections of online customer reviews representing digital cameras and vacuum cleaners. Concurrent with the automated processing, we also manually surveyed a smaller sample of reviews to confirm our intuitions about the distribution of need and attribute terms within reviews.

5.1 Needs and Attributes

Working from a smaller sample, two coders independently reviewed each of 401 reviews to extract phrases representing needs and attributes. The reviews ranged over 16 different

products from a minimum of $85.00 to a maximum of $1000.00. While many needs-phrases are relatively straightforward, *my children* or *do selling on eBay*, other phrases were considerably more vague. *400 pictures in Europe* was coded because it suggested *travel* and *8x10* was noted because it suggested *photo-enlarging* or *at-home printing*.

In all, the coders agreed on 972 distinct, needs-related phrases appearing a total of 1,942 times for an average of 3.3 needs per review. As expected, by contrast, the same reviews documented 2,382 distinct attribute-phrases appearing 5,252 times for an average of 13 attributes per review. The higher volume and the higher incidence of repetition among attribute tokens confirmed that rules relating needs to attributes would, at best, exhibit low support.

5.2 Digital Cameras

We began with a database of more than 8,000 digital camera customer reviews from Epinions.com. After parsing and filtering, the initial language processing resulted in 1,941 candidate need or attribute tokens. Application of the first stage, hyper-edge clustering, produced 46 non-disjoint clusters that exhibited very high confidence but relatively low support. After updating the review-token matrix with these clusters, the second stage of mining produced a set of low-frequency, high confidence hyper-rules that reduced the candidate set of needs tokens to 297. Although we did not do so, a post-processing stage might further clean the candidate set by combining synonyms, meronyms, hyponyms, etc. For example, *husband*, *wife*, and *kid* are all independently recorded. Likewise *trip* and *vacation* or *cost* and *price*.

We evaluated our candidate set of needs against needs drawn from a set of fifteen print and on-line buying guides summarized in Table 1. As noted in row 1, the guides list a minimum of five and a maximum of seventeen distinct needs. Consequently, the precision of our candidate set, containing 297 is small by comparison. We do observe that recall is consistently high. Moreover, recall penalties are due to two consistent omissions. Following standard marketing practice, the guides ask users to self-label themselves as *beginner*, *professional*, or various grades of *intermediate*, *amateur*, etc. While our analysis produced both the *beginner* and *professional* labels, it is worth noting that of the 8,000 digital camera reviews used, Noun Phrases containing *intermediate* or *amateur* were too infrequent to pass the frequency threshold for subsequent analysis. The second major omission was that of *landscape*. We simply were unable to elevate the support of a hyper-rule involving landscape photography high enough to include.

By contrast, it is worth noting the value that may come from inviting the voice of the consumer into the process of needs-based marketing or design. First, the same online guide(s) sometimes draw ambiguous distinctions between different uses. For example, *online photo album* v. *personal home page* v. *share pictures online*. Second, more significantly, the expert buying guides appeared to omit several needs categories that might otherwise be significant: Digital cameras as a replacement for SLR/film cameras, travel photography, and any number of variants on family, grandparents, grandchildren, or babies.

5.3 Vacuum Cleaners

We repeated the analysis on a set of 2,500 vacuum cleaner reviews, also from Epinions.com. The initial language pruning returned a candidate set of 611 tokens. The two-stage mining process then reduced that set to 123 tokens. As before, we compared our candidate needs to those taken from various print and on-line buying guides. Again, the precision is below any level of significance, but the recall perhaps highlights contrasts between the online user community and that of the buyer guides. We noted some ambiguity between *upholstery cleaning* and *furniture cleaning*, and likewise for *corners* v. *crevices*. We counted all as distinct terms for evaluation purposes. Our recall numbers were consistently penalized for missing *tile*, *upholstery*, and *drapes*. By contrast, the guides all appeared to focus on functional tasks, omitting higher-level need categories such as allergy, pets, and children.

Table 2. Comparing recall for needs.

	CR05	CR04	CR03	VACC	UNB	BVAC
P	7	6	5	2	8	22
R	5	5	2	1	7	12
	Lowes	Ebuy	Ivac	Prun	Skeg	
P	22	3	13	8	6	
R	10	3	7	5	3	

Though not in the guides, there are concepts that appear repeatedly throughout the reviews, lending a voice to the consumers' needs. Rather than a critique, the existence of a gap suggests an opportunity for managerial support by developing automated tools to process the continuous feeds of information enabled by the Internet.

6. CONCLUSION

In the present state of rapid development and short product lifecycles, automated processing of needs is a necessary precursor to decision support. However, relying upon publicly available text as a knowledge base provides the added advantage of transparency. The importance of explanation was established with the first expert systems; researchers quickly discovered that users were unwilling to accept automated support without the ability to view the underlying rationale [27]. Users need to know "why." By using reviews, we can literally link back to the explicit reviews that support an outcome.

Table 1. Comparing customer review recall for digital camera *needs* documented in several buying guides.

	Epin	Cnet	MP	PX	DCBG	Viewz	Buyz	Comp	PG	IA	AARP	CRDC	CR04	CR03	CR02
Orig	9	11	5	17	8	6	6	5	5	9	5	8	10	11	5
R	8	8	5	12	4	3	6	4	3	7	5	3	7	7	4

This paper describes work in progress; a number of limitations still require attention. Even with refinements, the scope of an automated solution would remain limited. As already discussed, while the method does promise some modest gains in acquiring customer needs, the resulting sets of candidate need tokens are still larger than might be pragmatic. In part, this could be addressed by a post-processing step to gather homonyms, meronyms, etc. [24].

There are also, however, some inherent limitations to the stage-one clustering based upon frequent item-sets. Because the graph encodes co-occurrences within individual reviews, the graph is less likely to encode synonyms, meronyms, etc. Indeed this intuition about review structure is exploited by Lee to learn the properties of product attributes [16]. Because hyperedges are less likely to encode such linguistic elements, hyper-rules, as currently defined, lose some of their semantics relating user needs to product attributes. A hyper-rule is supported by the Cartesian product of the power set of each constituent hyperedge. However, the mutual co-occurrence that defines a hyperedge, may have only limited bearing on the one or more vertices in the linked hyperedge. A cluster of synonyms or meronyms would both improve the underlying rule support and have obvious semantic consequence for helping to map the relationship between specific needs and attributes.

In addition, the hyper-edges formed by frequent item sets are not necessarily disjoint. If the clustering is applied repeatedly, subsequent rules between overlapping hyperedges could cycle. In the two-stage model, cycles should not occur provided the threshold for the rules is contained by the threshold for forming hyperedges. Rather than clustering based upon frequent item sets, then, we also wish to consider alternate graph clustering processes including the obvious relations to cliques in the dense structure.

Even with improved pre and post-processing, we appreciate that even continuous monitoring of customer reviews is no panacea for the problem of monitoring rapidly evolving customer needs. In learning from customer reviews, two frequent limitations are negation and fabrication. In negation, customers disclaim a need or attribute (e.g. "I did not buy this because ….") Fabricated comments, both positive and negative, can skew results much like packing a ballot box. More sophisticated text processing techniques exist for addressing negation and could be easily included here. Short of spam, we achieve scale by focusing on product-level analysis, aggregating all reviews for a product rather than focusing on a particular manufacturer or model. Interestingly enough, the government may also intervene. The European Union recently passed new consumer protection laws prohibiting false representations online [6].

Equally significant is the fact that customers are free to comment on any topic that they like within an online customer review. We make the strong underlying assumption that, although free to choose, customers are more likely to comment on those characteristics that are most relevant to their personal needs. This intuition breaks down in the face of comprehensive, professional reviews that attempt to cover all manner of users. We could pre-filter reviews by attempting to classify 'professional reviews' that are not likely to follow this model. For example, we can use review length, number of attributes (needs) mentioned, to separate out professional reviewers that may confuse the relationships. Alternatively, we could apply variations on topic mining approaches to segment documents into homogeneous regions: e.g. segmenting a review into comments on indoor photography and comments on outdoor photography.

From one-to-one marketing to user-centered design, the Internet has enabled marketers and manufacturers to draw closer to their customers. However, as evidenced in our comparison of print and online guides, a gap persists between customer needs and those who seek to meet those needs. We propose to augment existing techniques for studying customer needs through the automated, needs-based analysis of online customer reviews.

7. ACKNOWLEDGMENTS

This work was supported, in part, by grants from the University of Pennsylvania Research Foundation, the Fishman-Davidson Center for Service and Operations Management, and the Wharton eBusiness Initiative. The authors would also like to thank Esther Chen, Ellen Ngai, and Sojeong Song for their research assistance.

8. REFERENCES

[1] Adomavicius, G. and Tuzhilin, A. Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-art and Possible Extensions. IEEE Transactions of Knowledge and Data Engineering, 17 (6). 734-749.

[2] Borgelt, C. and Kruse, R., Induction of Association Rules: Apriori Implementation. in 15th Conf on Computational Statistics (Compstat), (Berlin, Germany, 2002).

[3] Chen, Y. and Xie, J. Online Customer Review: A New Element of Marketing Communications Mix SSRN, 2004.

[4] Cheung, K.-W., Kwok, J.T., Law, M.H. and Tsui, K.-C. Mining customer product ratings for personalized marketing. Decision Support Systems, 35. 231-243.

[5] Chevalier, J.A. and Mayzlin, D. The Effect of Word of Mouth on Sales: Online Book Reviews. Journal of Marketing Research, 43. 345-354.

[6] Coates, S. Consumer Protection Laws Times of London, London, UK, 2007.

[7] Dave, K., Lawrence, S. and Pennock, D.M., Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. in WWW, (2003), 519-528.

[8] Ghose, A., Ipeirotis, P. and Sundararajan, A. The Dimensions of Reputation in Electronic Markets, New York University, 2006, 32.

[9] Grenci, R.T. and Todd, P.A. Solutions-Driven Marketing. Communicatons of the ACM, 45 (3). 65-71.

[10] Han, E.-H., Karypis, G., Kumar, V. and Mobasher, B., Clustering Based on Association Rule Hypergraphs. in Workshop on Research Issues in Data Mining and Knowledge Discovery, (1997).

[11] Han, J. and Kamber, M. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, San Francisco, 2001.

[12] HOG. Harley-Davidson Owner's Group.

[13] Hu, M. and Liu, B., Mining and Summarizing Customer Reviews. in KDD04, (Seattle, WA, 2004), ACM.

[14] Kotler, P. Marketing Management. Prentice Hall, 2003.

[15] Lee, T., Use-centric mining of customer reviews. in Workshop on Information Technology and Systems (WITS), (Washington, D.C., 2004), 146-151.

[16] Lee, T.Y. Constraint-based Ontology Induction from Online Customer Reviews. Group Decision and Negotiation, 16. 255-281.

[17] Liu, B., Hu, M. and Cheng, J., Opinion Observer: Analyzing and Comparing Opinons on the Web. in WWW 2005, (Chiba, Japan, 2005), ACM.

[18] Liu, H. MontyLingua: An End-to-End Natural Language Processor with Common Sense, 2004.

[19] McEachern, T. and O'Keefe, B. Re-Wiring Business: Uniting Management and the Web. John WIley & Sons, Inc., New York, 1998.

[20] Nasukawa, T. and Yi, J., Sentiment Analysis: Capturing Favorability Using Natural Language Processing. in K-CAP`03, (Sanibel Island, Florida, 2003), ACM, 70-77.

[21] O'Keefe, R.H. and McEachern, T. Web-based Customer Decision Support Systems. Communicatons of the ACM, 41 (3). 71-78.

[22] OED. Oxford English Dictionary.

[23] Orman, L. Consumer Support Systems. Communicatons of the ACM, 50 (4). 49-54.

[24] Popescu, A.-M. and Etzioni, O., Extracting Product Features and Opinions from Reviews. in HLT-EMNLP, (2005).

[25] Popping, R. Computer-Assisted Text Analysis. Sage Publications, Thousand Oaks, CA, 2000.

[26] Randall, T., Terwiesch, C. and Ulrich, K.T. User Design of Customized Products. Marketing Science, 26 (2). 268-280.

[27] Rich, E. and Knight, K. Artificial Intelligence. McGraw-Hill, New York, 1991.

[28] Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H. and Jin, C., Red Opal: Product-Feature Scoring from Reviews. in ACM EC, (San Diego, CA, 2007).

[29] Skiena, S. The Algorithm Design Manual. TELOS, Springer-Verlag, New York, 1998.

[30] Stephens, C. Embracing Consumer Buzz Creates Measurement Challenges for Marketers Spark Research, Compete, Inc., Boston, 2006.

[31] Stolze, M. and Strobel, M. Recommending as Personalized Teaching. in Karat, C.-M., Blom, J.O. and Karat, J. eds. Designing Personalized User Experiences in eCommerce, Springer, Netherlands, 2004, 293-313.

[32] Toubia, O., Simester, D.I., Houser, J.R. and Dahan, E. Fast Polyhederal Adaptive Conjoint Estimation. Marketing Science, 22 (3). 273-303.

[33] Turney, P., Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. in, (2002), 417-424.

[34] Ulrich, K.T. and Eppinger, S.D. Product Design and Development. McGraw-Hill/Irwin, New York, 2003.

[35] Urban, G.L. and Hauser, J.R. "Listening In" to Find Unment Customer Needs and Solutions eBusiness@MIT, no. 156, MIT, Boston, MA, 2003, 36.

[36] Weber, B. Basic Content Analysis. Sage Publications, Newbury Park, CA, 1990.

[37] Yahoo and comScore. Engaging Advocates through Search and Social Media, 2006.